Proceedings of the 56th Hawaii International Conference on System Sciences | 2023

On Left and Right: Understanding the Discourse of Presidential Election in Social Media Communities

Lina Zhou
The University of North Carolina at Charlotte
lzhou8@uncc.edu

Jie Tao
Fairfield University
jtao@fairfield.edu

Kanlun Wang
The University of North Carolina at Charlotte
kwang17@uncc.edu

Abstract

As a promising platform for political discourse, social media becomes a battleground for presidential candidates as well as their supporters and opponents. Stance detection is one of the key tasks in the understanding of political discourse. However, existing methods are dominated by supervised techniques, which require labeled data. Previous work on stance detection is largely conducted at the post or user level. Despite that some studies have considered online political communities, they either only select a few communities or assume the stance coherence of these communities. Political party extraction has rarely been addressed explicitly. To address the limitations, we developed an unsupervised learning approach to political party extraction and stance detection from social media discourse. We also analyzed and compared (sub)communities with respect to their characteristics of political stances and parties. We further explored (sub)communities' shift in political stance after the 2020 US presidential election.

Keywords: political party, stance, presidential election, zero-shot learning, ensemble learning

1. Introduction

Social media becomes a promising platform for political discourse. Political candidates and their advocates use social media platforms to promote themselves and their policies and to attack their opponents and their policies (Darwish et al., 2017). The openness of social media platforms also encourages a broad participation of citizens who can express their opinions or state their positions, get feedback from peers, and understand diverse viewpoints on emerging issues (AlDayel & Magdy, 2021). Stance is the expression of standpoint and judgment toward a given proposition (Biber & Finegan, 1988), such as a political party, candidate, or policy. Since people including social media users can take different stances on a political matter, their discussion greatly enriches the political discourse. Therefore, social media can serve as a platform for collaboration and discussion on political matters by exposing its users to various political stances to better inform the users' political choices (Rathi et al., 2021) and promote depolarization (Wojcieszak et al., 2020). On the other hand, there has been a major concern about the role of social media in political polarization (Kubin & von Sikorski, 2021). Additionally, social media political advertising may evoke negative responses (Boerman & Kruikemeier, 2016). The differing effects of social media further highlight the importance of understanding the political discourse via the lens of social media communities.

The politics-related discussion on social media culminates around presidential elections in the U.S. For instance, social media was an integral part of the political campaign strategy of a former President Barack Obama for the 2008 election, which not only helped him increase campaign fundraising but also empowered volunteers (Aaker & Chang, 2009). Another former President Donald Trump utilized social media to drive attention (Bickart et al., 2017). The 2020 U.S. presidential election was no exception, engendering heated discussions on social media. Thus, we chose the presidential election as the context of this research. Existing studies of the presidential election in social media data have addressed the issues of election outcomes prediction, sentiment analysis, fake news detection, and the impacts of the presidential election. Stance detection can contribute to these issues because people tend to take positions on political issues. In view of the controversial nature of many political issues, the US presidential election invites debates and argumentation to help the public better understand the candidates' agendas and policies.

Prior work on stance detection has primarily used supervised learning techniques. Such techniques rely on labeled data which remains scarce and of varying quality. In addition, the analysis of stance in social media data has been mainly conducted at the user or post level (Al-Ghadir et al., 2021), which tends to require multiple and even a large number of posts from the same users,

URI: https://hdl.handle.net/10125/102677
978-0-9981331-6-4

and/or the information extracted from user profiles or interactions to perform well (e.g., Darwish et al., 2020; Dey et al., 2017; Ghosh et al., 2019; Li et al., 2021; Wei et al., 2018). Studies have approached other political issues in social media discourse from the perspective of (sub)communities, simply referred to as *sub-communities* hereafter. However, these studies assumed that individual contents from the same communities share the same political stance (e.g., De Francisci Morales et al., 2021; Jungherr et al., 2022). This does not reflect the complexity of real political discourse such as implicit expression of political stance and change of opinions. Further, like the traditional news sources that differ in partisan bias as left-leaning, center, or right-leaning (AllSides, 2019), there could be differences in the political stance across different social media platforms and even different sub-communities on the same platform. However, the latter has received much less attention in the stance detection literature.

Stances are expressed toward targets, such as presidential candidates and/or political parties that the candidates represent in the context of our research. In some cases, stances are explicitly linked to tailored targets, such as debates (Murakami & Raymond, 2010) and news (AllSides, 2019). However, in other cases such as online public discussions, complex content may involve one or more targets. Without heeding specific targets within content, the expressed stances can be obscure. Therefore, target extraction, and more specifically political party extraction, becomes a relevant and challenging research problem, which needs to address the variant and/or implicit references to the political parties in social media discourse. However, this problem has not been explicitly studied in the literature.

To address the above-mentioned limitations, this research is aimed at exploring the social media discourse pertaining to the U.S. 2020 presidential election via analyzing political stances and parties. Specifically, we answer the following research questions: Can we develop unsupervised learning techniques to extract political parties from social media discourse effectively? How about political stances? How homogeneous are sub-communities in terms of their expressions of political parties and political stances? How does the social media discourse of Republican Targeted sub-communities (RTC) compare with Democrat Targeted sub-communities (DTC)? Did sub-communities experience a shift in the overall political stance after the 2020 U.S. presidential election? We collected data from Reddit to answer these research questions.

This work makes multifold research contributions. First, we propose a zero-shot ensemble learning model to extract target political parties and detect political stances from social media discourse for the first time. Specifically, we enhance the zero-shot model by incorporating ensemble techniques such as meta-modeling that integrates multiple types of knowledge, including terms associated with latent topics extracted from the social media discourse and noisy labels generated by the zero-shot model. Second, the findings of this study in the context of the 2020 U.S. presidential election uncover the differences in stance diversity and post-election shift between social media communities targeting the two main political parties. Third, we provide evidence challenging the common assumption of stance coherence of online communities in the political domain. Last but not least, unlike previous studies on stance detection that are dominated by using tweets, we build stance detection models with Reddit posts. Compared with tweets, Reddit data presents unique challenges to stance detection partly because it does not use hashtags. In addition, we propose a multi-step approach to collecting relevant data from Reddit, including sorting-based post collection, sub-community selection and stance annotation, and sub-community-based data collection and filtering.

2. Background and related work

2.1. The role of social media in political discourse

There are two main political parties in the US: the Democratic Party and the Republican Party. They often take different political positions on the matter that can be characterized as left-wing vs. right-wing stances. Liberals generally vote Democrat and conservatives vote Republican, and thus liberal and conservative have also been used to refer to the ideologies of the two parties respectively. Social media has been used to promote political campaigns, share political opinions, discuss controversial issues, and predict election outcomes. Although there is a concern that social media may exacerbate political polarization partly due to selective exposure (Kubin & von Sikorski, 2021) and fake news and propaganda (Olaniran & Williams, 2020), there is counter-evidence of the depolarization effect owing to exposure to diverse information or different viewpoints (Wojcieszak et al., 2020). One study shows that following social media platforms and the activeness of political parties in using social media platforms influence political choice-making (Rathi et al., 2021). Another study finds that social media may influence certain groups' voting decisions (Fujiwara et al., 2020). In addition, there are significant differences in political polarization across different social media platforms (Yarchi et al., 2021). For instance, Facebook was found to be the most heterophilic platform in terms of political polarization, and depolarization was shaped over time on WhatsApp

(Yarchi et al., 2021). Thus, social media can serve as a great lens for examining political stances.

2.2. Stance detection in online discourse

Stance detection is a process of determining the authors' stance (supportive, opposing, or neutral) towards a target within the text (AlDayel & Magdy, 2021; Sobhani et al., 2017). In the case of political discourse, the target can be a presidential candidate, a policy, a proposition, and a social movement. In other words, stance detection requires two inputs (Li et al., 2021): a target that must be defined and an argument or comment that expresses a stance tendency toward the specific target.

In view of the large body of work on stance detection, we focus our review on studies of online discourse. Stance detection has been performed on data collected from online platforms, such as online discussion boards (Murakami & Raymond, 2010), Twitter (Mohammad et al., 2016), and websites (Murakami & Raymond, 2010). The SemEval 2016 stance detection task provides labeled Twitter data (Mohammad et al., 2016), triggering a growing body of research on stance detection. Unlike Twitter, which is a microblogging site, Reddit is classified as a social news aggregation site where contents are moderated and rated. The rating is based on different factors such as vote counts. Reddit is made up of user-created 'subreddits', which are sub-communities centered on specific topics such as political advocacy or opponent groups. In addition, each sub-community has its own goals and community norms which set the stage for user interactions. Given the unique characteristics of Reddit, it creates new opportunities for understanding the political stance.

2.3. Stance detection techniques

Existing machine learning techniques for stance detection can be categorized into three types (AlDayel & Magdy, 2021): supervised, weakly-supervised, and unsupervised learning techniques. First, supervised techniques can be further grouped into traditional classification techniques, such as Naive Bayes (NB), SVM (Dey et al., 2017; Wei et al., 2018), and KNN (Al-Ghadir et al., 2021), and deep learning techniques, such as RNNs (Benton & Dredze, 2018), Bi-LSTM and nested LSTMs (Siddiqua et al., 2019), LSTM + attention (Sun et al., 2018), bidirectional gated recurrent unit network (BiGRU) (Wei et al., 2018), and BERT (Li et al., 2021). The deep learning models consist of attention-based, convolution-based, word embedding-based, and transformer models (e.g., BERT) (Ghosh et al., 2019). The results of a comparative analysis of representative models in each category show that the selected transformer model achieved a significantly superior performance to other competing models due to the former model's ability to capture contextual information of the text (Ghosh et al., 2019). The supervised learning algorithms dominate existing methods for stance detection.

Weakly-supervised learning for stance detection can take many forms, such as leveraging a small amount of labeled data for different tasks, shared topics between the different targets, and label dependencies (AlDayel & Magdy, 2021; Benton & Dredze, 2018). For instance, the weakly supervised task in the SemEval 2016 task (Mohammad et al., 2016) provides a large number of tweets related to a single target but no training data.

Unsupervised learning techniques do not rely on labeled datasets. It can apply to the targets that have no or little training data. One study demonstrates a superior performance of unsupervised methods to traditional supervised counterparts (Darwish et al., 2020). However, the method works well only on highly active Twitter users, who tweet frequently on a target but perform poorly or very poorly on less active users. Additionally, they mainly use clustering algorithms, which do not reflect the state-of-the-art unsupervised techniques.

2.4. Zero-shot learning

Zero-shot learning is the latest development in the field of transfer learning, which treats supervised classification problems as unsupervised Natural Language Inference (NLI) tasks (Yin et al., 2019). Any NLI model takes as inputs sequence pairs (e.g., sentences, paragraphs), naming a premise and a hypothesis, and then determines whether the hypothesis supports, contradicts the premise, or neither. Zero-shot learning is usually used to learn noisy labels on vast amounts of data, since researchers believe that training models on large datasets with noisy labels are more generalizable, compared to traditional classification methods using accurately labeled but smaller datasets (Pushp & Srivastava, 2017). Specifically, zero-shot learning has been used in various Natural Language Processing (NLP) tasks, such as topic detection, emotion detection, and situational analysis.

Zero-shot learning treats the classification task (usually binary) as finding the relatedness between the texts and the classes. Although most zero-shot learning systems tend to be expensive to train, it is still considered more cost-effective compared to the extensive resources needed for data labeling. Additionally, leveraging pre-trained models (e.g., transformers) can largely lower the training costs. Because of the pre-training of the embedding and entailment models, it is recommended that both classes (i.e., words) and hypotheses (e.g., definitions, prompt texts) are used in zero-shot learning problems (Yin et al., 2019).

Zero-shot learning aims to scale classifying algorithms to discover new classes (label-partially-unseen) or transfer to new datasets (label-fully-unseen). In label-partially-unseen problems, different types of knowledge are used to assist the model decision-making process. For instance, Zhang et al., (2019) proposed a method combining ensemble learning on coarsely-grain labeled data, label hierarchy, and label-to-word paths to assist the zero-shot learning process. However, in the context of this study, the coarse labels (at the subreddit level) do not translate well to finer granularity (at the post level). In label-fully-unseen problems, the classifiers are unaware of the labels, and do not have any access to labeled data for task-specific training. In the context of this study, we employed models that are pre-trained on generic tasks such as NLI and datasets such as the Multi-Genre Natural Language Inference corpus (Williams et al., 2018), where the specific labels of political parties (e.g., democrat vs. republican) and stance (e.g., supportive vs. opposing) are unseen to the models. Thus, this study is categorized as a label-fully-unseen problem, which is more challenging but also more generalizable to the real world (Yin et al., 2019).

3. The proposed methods

We design two key artifacts: 1) methods for political party extraction and stance detection in the context of the US presidential election, and 2) a multi-step method for data collection and preparation.

3.1. Data collection and preparation

As discussed earlier, existing studies on stance detection in social media mainly used Twitter data. Despite a host of studies on using Reddit data to understand political discourse, they considered a few subreddit communities that were known to have distinct political stances. There is a separate stream of research on detecting rumor stances (Gorrell et al., 2019), but it does not explicitly address stance detection in the political discourse. To address the limited availability of labeled data, we propose a multi-step approach to collecting and preparing data from Reddit, which consists of three main steps: *sorting-based post collection*, *sub-community selection and stance annotation*, and *sub-community-based data collection and filtering*.

Sorting-based post collection. We first collected Reddit posts on the presidential election from the lists of top and relevance posts on a daily basis over the course of nine months around the Election Day (November 3, 2020) using the Reddit APIs (i.e., PRAW and Pushshift). The platform's top sorting mechanism shows posts with the highest votes from the most recent 24 hours, and its relevance sorting mechanism fetches posts that are most relevant to any user's search query. For each post, we collected its contents as well as the metadata such as subreddit community and post time.

Sub-community selection and stance annotation. After grouping posts by the sub-communities (843 total), we selected those sub-communities with five or more posts for stance annotation. The annotation task was focused on three types of information:

- political domain (whether the sub-community belongs to the political domain or not?)
- political party (is the sub-community targeted at Democratic, Republican, both or neither parties?)
- political stance (does the sub-community state a supportive, opposing, neutral, or no stance toward the target party?)

Two coders analyzed the selected sub-communities independently based on the latter's purpose statement and sample posts. The coders resolved the inconsistent results through discussion to reach a consensus. Based on the annotation results, we selected those sub-communities that belong to the political domain, target at Democratic, Republican party, or both; and take a stance toward the target party. As a result, 35 sub-communities met the inclusion criteria, among which 21 subreddits (e.g., r/JoeBiden and r/Conservative) explicitly stated their political stances in their purpose statements and thus were selected for data collection.

Sub-community-based data collection and filtering. We collected one-year worth of data from the selected sub-communities centering on the Election Day. We filtered the data using keywords related to the US presidential election, including political parties such as 'Democrat', 'Republican', 'GOP', and 'dems', and political candidates such as 'Joe', 'Biden', 'Donald', and 'Trump', resulting in 138,291 posts. We further filtered the data by removing the redundant, forbidden, bot-generated, and non-English posts, and posts with fewer than five words. The final dataset contains 9,717 posts (5,991 posts before and 3,726 posts after the presidential election).

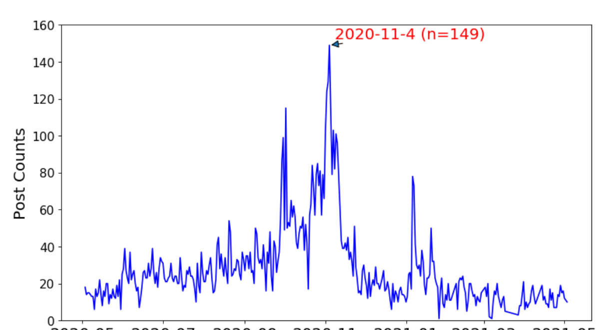

Figure 1. The temporal distribution of post counts

The distribution of post counts over the timeframe of data collection is shown in Figure 1. As expected, the volume of posts peaked around the Election Day. There are also two other noticeable peaks. One occurred before the election (i.e., October 2^{nd}, 2020) when former President Trump and the first lady tested positive for COVID-19, and the other post-election (January 6^{th} - 7^{th}, 2021) when the capital riot took place.

3.2. Stance detection and party extraction

Both stance detection and party extraction can be considered as a multi-label, multi-class classification problem. We propose an ensemble-based method utilizing zero-shot learning, namely **E**nsemble **Z**ero-Shot **Class**ifier (**EZClass**), to extract political parties and detect stances from social media discourse. The architecture of the member models utilizing zero-shot learning is shown in Figure 2, whereas the ensemble methods are illustrated in Figure 3.

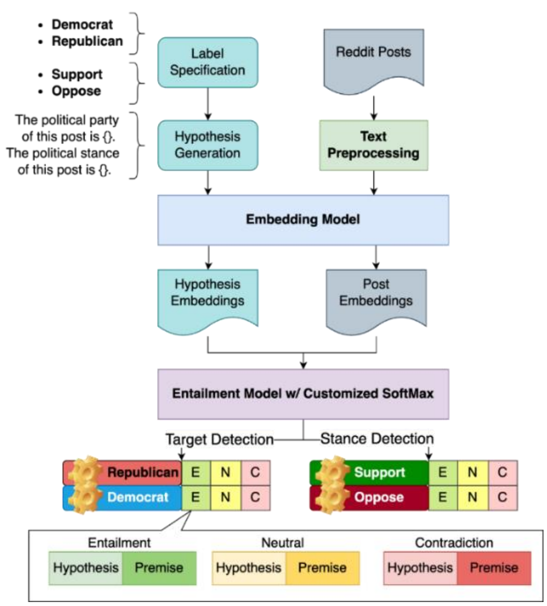

Figure 2. The zero-shot learning model

3.2.1. Zero-shot learning

To formulate text classification as a zero-shot learning problem, the texts to be classified are treated as premises and the classes as hypotheses. For instance, in the context of party detection, a sample premise is "*as a long-time Biden supporter, I would consider voting for Trump in this election*"; and the hypothesis is that *the political party of this post is Republican/Democrat*. For each of the hypotheses, the method assesses it against the three relationships, namely entailment (E), contradiction (C), and neutral (N)) by calculating the corresponding probabilities. The entailment process is to imitate how humans decide the true class(es) from any aspect of a text (Yin et al., 2019). In the case of single-label classification, the probabilities of the entailment relationship are fed into a softmax function to determine the most favorable class; while in multi-label classification, the probabilities of all three relationships are calculated for each label independently, to select all identified labels. In this study, we adopted the multi-label entailment model since our problem is to determine both party and stance of each post. We first conducted a pilot study with a small subset of the data to determine the labels and hypotheses for the entailment model. Based on the results of the pilot study, we selected party labels as "democrat", "Biden", "republican", and "Trump", because they yielded the best performance, with the hypothesis template of "The political party of the post is {}". For stance detection, we selected "supportive" and "oppose" as labels, along with "The political stance of the post is {}" as the hypothesis template.

The entailment model requires the premises and the hypotheses to be represented in the same embedding space. Transformers, as the latest development in NLP, are widely used to learn text embeddings. Given that zero-shot learning is an NLI task, it is reasonable to select the transformer models that are pre-trained for NLI tasks. Thus, we adopted the seven most popular models (based on a popular model repository (Models - Hugging Face, n.d.). Before feeding the posts to the embedding models, we performed standard text preprocessing tasks, including removing URLs, phone numbers, numbers, whitespace, hyphenated words, and quotation marks, as well as normalizing text encoding.

The pre-processed texts are then fed into the member models. To maintain the generalizability of the proposed method, we employed small-, medium-, and large-sized pre-trained models as the member models. We found that larger models tend to be better at detecting political parties from text content. In addition, the original softmax function appeared to be insufficient for both party and stance classifications using the entailment model. For instance, party extraction is a multi-label classification problem, as discussed earlier. Based on the results of our pilot study, including class labels such as "both" and "neither" can heavily impact model outputs. As a result, we chose to use initial labels (Democrat vs. Republican) and proposed a customized softmax function to accommodate the multi-label classification. First, we summed the probabilities of "democrat" and "Biden" as well as "republican" and "Trump" as the normalized classification probabilities, respectively. If only one of the probabilities (Democrat/Biden vs. Republican/Trump) from the original softmax function was greater than a given threshold (e.g., 0.5), then the associated label would be selected; if both probabilities were greater than the given threshold, then post would be classified as "both"; otherwise the post would be classified as "neither". For stance detection, we customize the softmax function in a similar way. If probabilities of both supportive and opposing stances were either greater than or less than the threshold, the post would be classified as "neutral".

3.2.2. Ensemble and knowledge enhancement

Although our tweaking of the premises and hypotheses and customized softmax function enhanced the model performance, the individual learning models are considered "weak learners" due to the nature of the NLI tasks (Zhang et al., 2021). Therefore, ensemble learning has the potential to boost classification performance. In this study, we designed two types of ensemble learning models: majority vote and knowledge-enhanced meta-modeling (see Figure 3). Besides the internal knowledge enhancement in zero-shot learning (with noisy labels from party extraction to stance detection), we propose a topic modeling-based knowledge enhancement method.

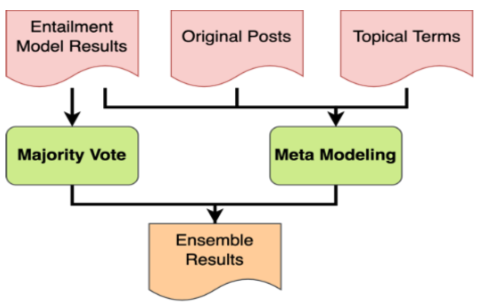

Figure 3. Ensemble learning and knowledge enhancements

Voting-based methods (bagging) are widely used to boost the performance of classifier ensembles by reducing the variance of the results across the member models. We selected the majority vote in this study since the results from our pilot study indicated weighted voting (using the entailment probabilities) did not show a significant difference in performance, but a significant difference in computational complexity.

One way for knowledge enhancement is to use the noisy labels from zero-shot learning for both classification tasks in EZClass (e.g., party labels for party extraction, and both party/stance labels for stance detection) before performing zero-shot classification. The rationale for such a design decision is that party labels are helpful for stance detection since the latter is dependent on the former. Compared to the traditional meta-modeling methods, which only utilize the outputs from the member models (i.e., political party and stance labels), our model is able to learn from the post contents as well. To this end, we employed the bart-large-mnli model (Lewis et al., 2019) as the meta-model.

Compared to the majority vote, the meta-modeling technique is suggested as the more appropriate ensemble method for zero-shot text classification problems (Puri & Catanzaro, 2019). In addition, a voting-based ensemble requires all member models to perform (relatively) well on the classification task, which however is hard to determine in zero-shot learning. On the other hand, meta-modeling ensembles (stacking) can relax the above assumption on modeling results. To incorporate domain knowledge into the meta-models, we first append the class labels (e.g., political party and/or stance) to the post texts.

To further enhance knowledge fusion in EZClass, we also decided to append topical terms to the original posts. To extract the topical terms, we leveraged a transformer-based topic modeling approach, BERTopic (Grootendorst, 2022), to extract topics from the posts. Compared to traditional topic modeling techniques, such as Latent Dirichlet Allocation or Non-negative Matrix, BERTopic is more context-dependent, position-aware, representative beyond the word level, and better at handling out-of-vocabulary words. We extracted the topic with the highest probability for each post, and then selected the top-N (e.g., 5) words associated with the topic (ranked by the relevance score) and appended them to the original posts. Finally, we combined the two meta-modeling strategies by appending both labels and topical terms to the original posts.

3.3. Sub-community and comparative analysis

To address the research questions about sub-communities, we used stance diversity and party diversity as variables. The diversity of sub-community s, $D(s)$, was defined as entropy (see eq. (1)), where $p(c)$ denotes the probability of class c in all posts from s, and P and S denote the set of political parties or stances, respectively. To operationalize the diversity measure, we leveraged the output classes of the best models for political party extraction and stance detection (see Section 3.2).

$$D(s) = -\sum_{c \in P,S} p(c) \, log_2 \, p(c) \qquad (1)$$

To compare the different types of sub-communities, we used the focal political party of a sub-community (DTC vs. RTC) as the independent variable and stance diversity and party diversity as the dependent variables respectively, and performed independent-samples t-tests. We repeated the above test by using the political stance of sub-communities (supportive vs. opposing) as the independent variable.

To address the research questions about stance shift, we compared the distributions of political party and stance in sub-communities between the pre-and post-election periods using the Chi-squared test. In addition, we performed this test for RTC and DTC separately.

4. Results and discussion

4.1. Party and stance detection

We prepared datasets in multiple steps to support the evaluation of our proposed models. First, we conducted a pilot study by randomly selecting a subset of 258 posts from the sorting-based post collection (see Section 3.1). Two researchers manually reviewed those posts to determine the expressed political party and stance independently. The inconsistent coding results were resolved via group discussion consisting of the two coders and a third researcher as an adjudicator. Based on the analysis results and our experience with the pilot study, we developed instructions for data annotation. Then, we randomly sampled 50 posts from the final subreddit data collection (see Section 3.1) to validate our labeling instruction. Finally, we selected another random set of 200 posts for manual annotation, and the resulting dataset was used for our formal evaluations.

We adopted the seven most popular models based on a popular model repository (Models - Hugging Face, n.d.). In addition, we performed ablation experiments by incrementally incorporating different components of our proposed methods, including majority vote and knowledge-enhanced meta-modeling. The latter further consists of three different settings, posts+labels (appending party labels to posts), post+terms (appending terms to posts), and post+terms+labels.

Given the imbalanced stance distribution in our dataset, we selected the weighted average F1 score (see eq. (2)) as the evaluation metric, where the F1 score is a harmonic mean of precision and recall. The F1-scores of individual classes ($F1(y_c, \hat{y}_c),\ c \in C$) are derived as the average of the support of y_c ($|y_c| / \sum_{c \in C} |y_c|$), where $y_c \subseteq y$ is from class c, and $\hat{y}_c$ is the predicted class labels for y_c.

$$Weighted\ F1 = \frac{1}{\sum_{c \in C} |y_c|} \sum_{c \in C} |y_c|\, F1(y_c, \hat{y}_c) \quad (2)$$

The model performances are reported in Table 1. The best result for each setting is highlighted in bold. The results of party extraction show that the post+terms model outperforms other alternative models with a weighted F1 score of 0.8. For the detection of political stance, posts+labels yield the best performance, with a weighted F1 score of 0.55.

To further validate the model performances in stance detection, we compared the overall stances between supportive and opposing sub-communities based on the stances of their individual posts. The overall stance of each sub-community was measured as the ratio of positive to negative stance in its posts. The results of the independent-sample t-test show that the overall stance of supportive communities (mean=.589) is significantly higher (p<.01) than that of opposing counterparts (mean =.218). Similarly, we validated the performances in party extraction by comparing the overall targeted parties between DTC and RTC based on the information extracted from their individual posts. The overall targeted party was measured as the ratio of posts targeting democrats to those targeting republicans. The results show that the value of DTC (mean=1.390) is significantly higher (p<.001) than that of RTC (mean =.225). The above results validate the effectiveness of our proposed models.

Table 1. Model performances on party extraction and stance detection

EZClass Models	Political Party	Political Stance
Majority Vote	0.76	0.52
Post+Labels	0.73	**0.55**
Post+Terms	**0.80**	0.47
Post+Terms+Labels	0.78	0.46

4.2. Sub-community analysis

The descriptive statistics of the party diversity and stance diversity are reported in Table 2. Based on the outputs of party extraction, we plotted the party diversity and distribution in Figure 4, which were sorted in ascending order of party diversity. The figure shows that party diversity varies widely across different sub-communities, with the maximum value nearly doubling the minimum value. It can also be observed that the low diversity end (e.g., bottom 5) is dominated by RTC whereas the high diversity end (e.g., top 7) is mostly DTC except for r/AskThe_Donald. These observations suggest that the RTC is more likely to either target a single party or no party, whereas DTC is more likely to either target both parties or not mention any party than targeting a single party.

Table 2. Descriptive statistics (mean[sd]) of party and stance diversity

Target Party	Party Diversity	Stance Diversity
Democrat	1.45 [.156]	1.31 [.177]
Republican	1.20 [.204]	1.32 [.214]

We produce a plot of stance diversity and distribution based on the stance detection outputs in Figure 5. The figure shows that those sub-communities ranked toward the high-diversity end (e.g., top-5) are all RTC. They imply that some RTC tend to express mixed stances that both promote their own viewpoints and criticize those of the opposing parties. However, the sub-communities with low stance diversity consist of a mixture of DTC and RTC. In addition, most of the sub-communities, particularly those with low stance diversity, are dominated by the opposing stance, which expresses criticism of the target parties of their interests.

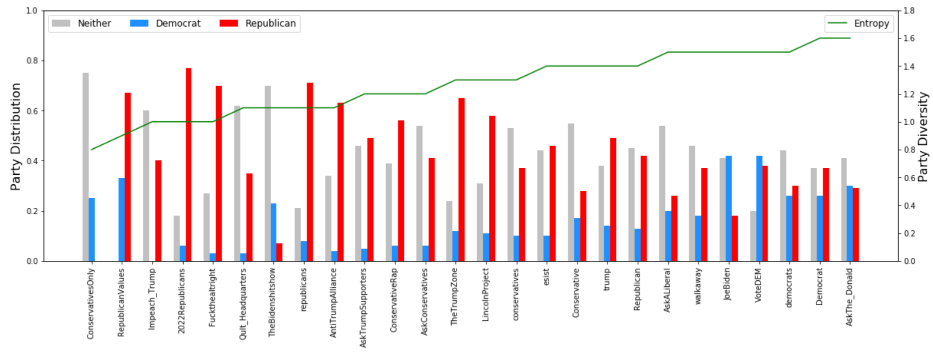

Figure 4. Political party distribution and party diversity of sub-communities

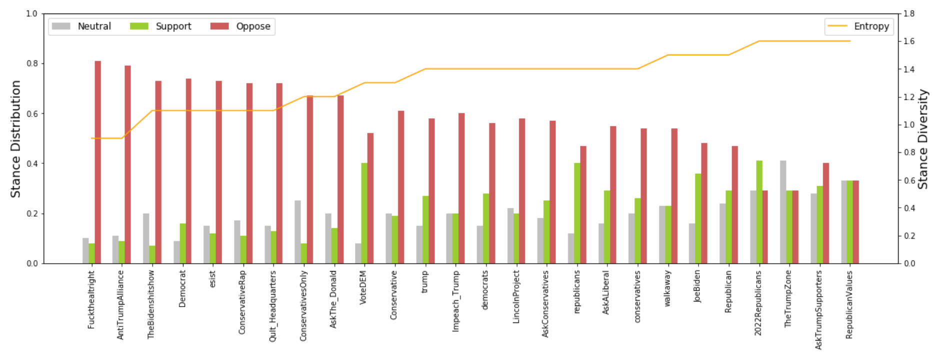

Figure 5. Political stance distribution and stance diversity of sub-communities

4.3. Comparative analyses

The t-test results of comparing party diversity between DTC and RTC reveal a significant difference (p<.01), with DTC showing a higher level of party diversity than RTC. The analysis of stance diversity did not yield any significant difference between DTC and RTC (p>.05).

The analysis results of stance diversity between sub-communities with different political stances show that the sub-communities with a supportive stance (mean=1.39, std = .155) have a higher stance diversity than those with an opposing stance (p<.01; mean=1.17, std=.217). They suggest that the advocacy sub-communities tend to use mixed stances, whereas opposition sub-communities tend to adopt a consistent stance.

The Chi-square test results show that stance distribution in DTC shifted significantly (p<.05) after the presidential election, while there was only a slight change in stance distribution for RTC (p<.1). Specifically, there was an increase in the percentage of positive stances in DTC, and a slight drop in the percentage of negative stances in RTC following the election.

5. Conclusion and future work

The main findings and research implications of this study are summarized as follows. First, using zero-shot and ensemble learning techniques, we built models to detect political stances and extract target parties from social media posts. The empirical results demonstrate that our proposed model — EZClass yields reasonable performance for both party extraction and stance detection. In addition, knowledge enhancement contributes to model performance. Second, our analysis of social media discourse at the sub-community level reveals within-community variances with respect to political parties and stances. Such findings challenge the assumption that individual contents in a sub-community are uniformly aligned with the stated political purposes of the sub-community. Third, the results of our comparative analyses uncover some interesting cross-community differences with respect to the characteristics of their political party and stance, and shifts in political stance following the presidential election.

The findings of this study have practical implications for social media platforms, community moderators and members, and political stakeholders. Social media platforms may leverage EZClass to facilitate stance categorization and summarization (Al-Ghadir et al., 2021) and even identify relevant arguments for different stances. In addition, community moderators may use our proposed models to support content moderation to enforce community norms and rules. The models can also help community members reduce information overload through the detection of political parties and stances. Further, the deep understanding of the complex political discourse this study provides can help political stakeholders formulate effective strategies for using social media to achieve their political goals.

This work has several limitations, which motivate future research issues. First, the data used in the model evaluation was relatively small due to its reliance on manual annotation. It would be desirable to validate our proposed models with a larger dataset. Second, we adopted the seven most popular pre-trained models in learning text embeddings. Another promising direction is to incorporate a self-training strategy using reinforcement learning into zero-shot learning (Ye et al., 2020). In addition, semi-supervised learning (e.g., combining zero-shot learning with traditional classification) may help boost the performance of stance detection models. Third, we used the post content and lexical terms to improve model performance in stance detection and party extraction. Our model architecture can be extended to incorporate other forms of knowledge such as knowledge graphs (Chen et al., 2022). Fourth, we treated different political parties as independent to generate model outputs. However, a single social media content may express political stances toward multiple parties. In other words, neither political stance nor party categories are mutually exclusive. One possible solution is to develop multi-label stance detection models (Sobhani et al., 2017), which can leverage the information about the relationships between different targets.

References

Aaker, J., & Chang, V. (2009). *Obama and the Power of Social Media and Technology | Stanford Graduate School of Business*. https://www.gsb.stanford.edu/faculty-research/case-studies/obama-power-social-media-technology

AlDayel, A., & Magdy, W. (2021). Stance detection on social media: State of the art and trends. *Information Processing and Management: An International Journal*, *58*(4). https://doi.org/10.1016/j.ipm.2021.102597

Al-Ghadir, A. I., Azmi, A. M., & Hussain, A. (2021). A novel approach to stance detection in social media tweets by fusing ranked lists and sentiments. *Information Fusion*, *67*, 29–40. https://doi.org/10.1016/j.inffus.2020.10.003

AllSides. (2019, May 27). AllSides. https://www.allsides.com/unbiased-balanced-news

Benton, A., & Dredze, M. (2018). Using Author Embeddings to Improve Tweet Stance Classification. *Proceedings of the 2018 EMNLP Workshop W-NUT: The 4th Workshop on Noisy User-Generated Text*, 184–194. https://doi.org/10.18653/v1/W18-6124

Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse Processes*, *11*(1), 1–34. https://doi.org/10.1080/01638538809544689

Bickart, B., Fournier, S., & Nisenholtz, M. (2017, March 1). What Trump Understands About Using Social Media to Drive Attention. *Harvard Business Review*. https://hbr.org/2017/03/what-trump-understands-about-using-social-media-to-drive-attention

Boerman, S. C., & Kruikemeier, S. (2016). Consumer responses to promoted tweets sent by brands and political parties. *Computers in Human Behavior*, *65*, 285–294. https://doi.org/10.1016/j.chb.2016.08.033

Chen, Q., Wang, W., Huang, K., & Coenen, F. (2022). Zero-Shot Text Classification via Knowledge Graph Embedding for Social Media Data. *IEEE Internet of Things Journal*, *9*(12), 9205–9213. https://doi.org/10.1109/JIOT.2021.3093065

Darwish, K., Magdy, W., & Zanouda, T. (2017). Trump vs. Hillary: What Went Viral During the 2016 US Presidential Election. In G. L. Ciampaglia, A. Mashhadi, & T. Yasseri (Eds.), *Social Informatics* (pp. 143–161). Springer International Publishing. https://doi.org/10.1007/978-3-319-67217-5_10

Darwish, K., Stefanov, P., Aupetit, M., & Nakov, P. (2020). *Unsupervised User Stance Detection on Twitter* (arXiv:1904.02000). arXiv. https://doi.org/10.48550/arXiv.1904.02000

De Francisci Morales, G., Monti, C., & Starnini, M. (2021). No echo in the chambers of political interactions on Reddit. *Scientific Reports*, *11*(1), 2818. https://doi.org/10.1038/s41598-021-81531-x

Dey, K., Shrivastava, R., & Kaushik, S. (2017). Twitter Stance Detection—A Subjectivity and Sentiment Polarity Inspired Two-Phase Approach. *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*, 365–372. https://doi.org/10.1109/ICDMW.2017.53

Fujiwara, T., Müller, K., & Schwarz, C. (2020). *The Effect of Social Media on Elections: Evidence from the United States* (SSRN Scholarly Paper No. 3719998). Social Science Research Network. https://doi.org/10.2139/ssrn.3719998

Ghosh, S., Singhania, P., Singh, S., Rudra, K., & Ghosh, S. (2019). Stance Detection in Web and Social Media: A Comparative Study. In F. Crestani, M. Braschler, J. Savoy, A. Rauber, H. Müller, D. E. Losada, G. Heinatz Bürki, L. Cappellato, & N. Ferro (Eds.), *Experimental IR Meets Multilinguality, Multimodality, and Interaction* (pp. 75–87). Springer International Publishing. https://doi.org/10.1007/978-3-030-28577-7_4

Gorrell, G., Kochkina, E., Liakata, M., Aker, A., Zubiaga, A., Bontcheva, K., & Derczynski, L. (2019). SemEval-2019 Task 7: RumourEval, Determining Rumour

Veracity and Support for Rumours. *Proceedings of the 13th International Workshop on Semantic Evaluation*, 845–854. https://doi.org/10.18653/v1/S19-2147

Grootendorst, M. (2022). *BERTopic: Neural topic modeling with a class-based TF-IDF procedure* (arXiv:2203.05794). arXiv. https://doi.org/10.48550/arXiv.2203.05794

Jungherr, A., Posegga, O., & An, J. (2022). Populist Supporters on Reddit: A Comparison of Content and Behavioral Patterns Within Publics of Supporters of Donald Trump and Hillary Clinton. *Social Science Computer Review*, *40*(3), 809–830. https://doi.org/10.1177/0894439321996130

Kubin, E., & von Sikorski, C. (2021). The role of (social) media in political polarization: A systematic review. *Annals of the International Communication Association*, *45*(3), 188–206. https://doi.org/10.1080/23808985.2021.1976070

Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). *BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension* (arXiv:1910.13461). arXiv. https://doi.org/10.48550/arXiv.1910.13461

Li, Y., Sun, Y., & Zhu, N. (2021). BERTtoCNN: Similarity-preserving enhanced knowledge distillation for stance detection. *PLOS ONE*, *16*(9), e0257130. https://doi.org/10.1371/journal.pone.0257130

Models—Hugging Face. (n.d.). Retrieved June 15, 2022, from https://huggingface.co/models

Mohammad, S., Kiritchenko, S., Sobhani, P., Zhu, X., & Cherry, C. (2016). SemEval-2016 Task 6: Detecting Stance in Tweets. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 31–41. https://doi.org/10.18653/v1/S16-1003

Murakami, A., & Raymond, R. (2010). Support or Oppose? Classifying Positions in Online Debates from Reply Activities and Opinion Expressions. *Coling 2010: Posters*, 869–875. https://aclanthology.org/C10-2100

Olaniran, B., & Williams, I. (2020). Social Media Effects: Hijacking Democracy and Civility in Civic Engagement. In J. Jones & M. Trice (Eds.), *Platforms, Protests, and the Challenge of Networked Democracy* (pp. 77–94). Springer International Publishing. https://doi.org/10.1007/978-3-030-36525-7_5

Puri, R., & Catanzaro, B. (2019). *Zero-shot Text Classification With Generative Language Models* (arXiv:1912.10165). arXiv. https://doi.org/10.48550/arXiv.1912.10165

Pushp, P. K., & Srivastava, M. M. (2017). *Train Once, Test Anywhere: Zero-Shot Learning for Text Classification* (arXiv:1712.05972). arXiv. https://doi.org/10.48550/arXiv.1712.05972

Rathi, M., Ghosh, A., Kumari, S., Sarkar, A., & Das, A. (2021). Influence of social media on the Political Choice Making: An Exploratory Study. *Turkish Online Journal of Qualitative Inquiry*, *12*(6), 8326–8339.

Siddiqua, U. A., Chy, A. N., & Aono, M. (2019). Tweet Stance Detection Using an Attention based Neural Ensemble Model. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 1868–1873. https://doi.org/10.18653/v1/N19-1185

Sobhani, P., Inkpen, D., & Zhu, X. (2017). A Dataset for Multi-Target Stance Detection. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, 551–557. https://aclanthology.org/E17-2088

Sun, Q., Wang, Z., Zhu, Q., & Zhou, G. (2018). Stance Detection with Hierarchical Attention Network. *Proceedings of the 27th International Conference on Computational Linguistics*, 2399–2409. https://aclanthology.org/C18-1203

Wei, P., Mao, W., & Zeng, D. (2018). A Target-Guided Neural Memory Model for Stance Detection in Twitter. *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–8. https://doi.org/10.1109/IJCNN.2018.8489665

Williams, A., Nangia, N., & Bowman, S. R. (2018). *A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference* (arXiv:1704.05426). arXiv. https://doi.org/10.48550/arXiv.1704.05426

Wojcieszak, M., Winter, S., & Yu, X. (2020). Social Norms and Selectivity: Effects of Norms of Open-Mindedness on Content Selection and Affective Polarization. *Mass Communication and Society*, *23*(4), 455–483. https://doi.org/10.1080/15205436.2020.1714663

Yarchi, M., Baden, C., & Kligler-Vilenchik, N. (2021). Political Polarization on the Digital Sphere: A Cross-platform, Over-time Analysis of Interactional, Positional, and Affective Polarization on Social Media. *Political Communication*, *38*(1–2), 98–139. https://doi.org/10.1080/10584609.2020.1785067

Ye, Z., Geng, Y., Chen, J., Chen, J., Xu, X., Zheng, S., Wang, F., Zhang, J., & Chen, H. (2020). Zero-shot Text Classification via Reinforced Self-training. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 3014–3024. https://doi.org/10.18653/v1/2020.acl-main.272

Yin, W., Hay, J., & Roth, D. (2019). *Benchmarking Zero-shot Text Classification: Datasets, Evaluation and Entailment Approach* (arXiv:1909.00161). arXiv. https://doi.org/10.48550/arXiv.1909.00161

Zhang, J., Lertvittayakumjorn, P., & Guo, Y. (2019). *Integrating Semantic Knowledge to Tackle Zero-shot Text Classification* (arXiv:1903.12626). arXiv. https://doi.org/10.48550/arXiv.1903.12626

Zhang, J., Yu, Y., Li, Y., Wang, Y., Yang, Y., Yang, M., & Ratner, A. (2021). *WRENCH: A Comprehensive Benchmark for Weak Supervision* (arXiv:2109.11377). arXiv. https://doi.org/10.48550/arXiv.2109.11377